18.175 Theory of Probability
Fall 2008

Section 23

Donsker Invariance Principle.

In this section we show how the Brownian motion W_t arises in a classical central limit theorem on the space of continuous functions on $\mathbb{R}_+$. When working with continuous processes defined on $\mathbb{R}_+$, such as the Brownian motion, the metric $\|\cdot\|_\infty$ on $C(\mathbb{R}_+)$ is too strong. A more appropriate metric d can be defined by

$$d_n(f,g) = \sup_{0\le t\le n} |f(t)-g(t)| \quad \text{and} \quad d(f,g) = \sum_{n\ge 1} \frac{1}{2^n} \frac{d_n(f,g)}{1+d_n(f,g)}.$$

It is obvious that $d(f_j,f)\to 0$ if and only if $d_n(f_j,f)\to 0$ for all $n\ge 1$, i.e. d metrizes uniform convergence on compacts. $(C(\mathbb{R}_+),d)$ is also a complete separable space, since any sequence is Cauchy in d if and only if it is Cauchy for each d_n. When proving uniform tightness of laws on $(C(\mathbb{R}_+),d)$, we will need a characterization of compacts via the Arzela-Ascoli theorem, which in this case can be formulated as follows. For a function $x\in C[0,T]$, its modulus of continuity is defined by

$$m^T(x,\delta) = \sup\Bigl\{|x_a - x_b| : |a-b|\le\delta,\ a,b\in[0,T]\Bigr\}.$$

Theorem 55 *(Arzela-Ascoli) A set K is compact in $(C(\mathbb{R}_+),d)$ if and only if K is closed, uniformly bounded and equicontinuous on each interval $[0,n]$. In other words,*

$$\sup_{x\in K}|x_0| < \infty \quad \textit{and} \quad \lim_{\delta\to 0}\sup_{x\in K} m^T(x,\delta) = 0 \quad \textit{for all} \quad T>0.$$

Here is the main result about the uniform tightness of laws on $(C(\mathbb{R}_+),d)$, which is simply a translation of the Arzela-Ascoli theorem into probabilistic language.

Theorem 56 *The sequence of laws $(\mathbb{P}_n)_{n\ge 1}$ on $(C(\mathbb{R}_+),d)$ is uniformly tight if and only if*

$$\lim_{\lambda\to+\infty}\sup_{n\ge 1}\mathbb{P}_n(|x_0|>\lambda) = 0 \tag{23.0.1}$$

and

$$\lim_{\delta\downarrow 0}\sup_{n\ge 1}\mathbb{P}_n(m^T(x,\delta)>\varepsilon) = 0 \tag{23.0.2}$$

for any $T>0$ and any $\varepsilon>0$.

Proof. $\Longrightarrow$. For any $\gamma>0$, there exists a compact K such that $\mathbb{P}_n(K) > 1-\gamma$ for all $n\ge 1$. By the Arzela-Ascoli theorem, $|x_0|\le\lambda$ for some $\lambda>0$ and for all $x\in K$ and, therefore,

$$\sup_n \mathbb{P}_n(|x_0|>\lambda) \le \sup_n \mathbb{P}_n(K^c) \le \gamma.$$

Also, by equicontinuity, for any $\varepsilon>0$ there exists $\delta_0>0$ such that for $\delta<\delta_0$ and for all $x\in K$ we have $m^T(x,\delta)<\varepsilon$. Therefore,

$$\sup_n \mathbb{P}_n(m^T(x,\delta)>\varepsilon) \le \sup_n \mathbb{P}_n(K^c) \le \gamma.$$

$\Longleftarrow$. Given $\gamma > 0$, find $\lambda_T > 0$ such that

$$\sup_n \mathbb{P}_n(|x_0| > \lambda_T) \le \frac{\gamma}{2^{T+1}}.$$

For each $k \ge 1$, find $\delta_k > 0$ such that

$$\sup_n \mathbb{P}_n\Bigl(m^T(x, \delta_k) > \frac{1}{k}\Bigr) \le \frac{\gamma}{2^{T+k+1}}.$$

Define a set

$$A_T = \Bigl\{x : |x_0| \le \lambda_T, m^T(x, \delta_k) \le \frac{1}{k} \text{ for all } k \ge 1\Bigr\}.$$

Then for all $n \ge 1$,

$$\mathbb{P}_n(A_T) \ge 1 - \frac{\gamma}{2^{T+1}} - \sum_k \frac{\gamma}{2^{T+k+1}} = 1 - \frac{\gamma}{2^T}.$$

By the Arzela-Ascoli theorem, the set $A = \bigcap_{T \ge 1} A_T$ is compact on $(C(\mathbb{R}_+), d)$ and for all $n \ge 1$,

$$\mathbb{P}_n(A) \ge 1 - \sum_{T \ge 1} \frac{\gamma}{2^T} = 1 - \gamma.$$

This proves that the sequence $(\mathbb{P}_n)$ is uniformly tight.

$\square$

Of course, for the uniform tightness on $(C[0,1], \|\cdot\|_\infty)$ we only need the second condition (23.0.2) for $T = 1$. Also, it will be convenient to slightly relax (23.0.2) and replace it with

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}_n(m^T(x, \delta) > \varepsilon) = 0. \tag{23.0.3}$$

Indeed, given $\gamma > 0$, find δ_0, n_0 such that for $\delta < \delta_0$ and $n > n_0$,

$$\mathbb{P}_n(m^T(x, \delta) > \varepsilon) < \gamma.$$

For each $n \le n_0$ we can find δ_n such that for $\delta < \delta_n$,

$$\mathbb{P}_n(m^T(x, \delta) > \varepsilon) < \gamma$$

because $m^T(x, \delta) \to 0$ as $\delta \to 0$ for all $x \in C(\mathbb{R}_+)$. Therefore,

$$\text{if} \quad \delta < \min(\delta_0, \delta_1, \ldots, \delta_{n_0}) \quad \text{then} \quad \mathbb{P}_n(m^T(x, \delta) > \varepsilon) < \gamma$$

for all $n \ge 1$.

Donsker invariance principle. Let us now give a classical example of convergence on $(C(\mathbb{R}_+), d)$ to the Brownian motion W_t. Consider a sequence $(X_i)_{i \ge 1}$ of i.i.d. random variables such that $\mathbb{E}X_i = 0$ and $\sigma^2 = \mathbb{E}X_i^2 < \infty$. Let us consider a continuous partial sum process on $[0, \infty)$ defined by

$$W_t^n = \frac{1}{\sqrt{n}\sigma} \sum_{i \le \lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor) \frac{X_{\lfloor nt \rfloor + 1}}{\sqrt{n}\sigma},$$

where $\lfloor nt \rfloor$ is the integer part of nt, $\lfloor nt \rfloor \le nt < \lfloor nt \rfloor + 1$. Since the last term in W_t^n is of order $n^{-1/2}$, for simplicity of notations, we will simply write

$$W_t^n = \frac{1}{\sqrt{n}\sigma} \sum_{i \le nt} X_i$$

and treat nt as an integer. By the central limit theorem,

$$\frac{1}{\sqrt{n}\sigma} \sum_{i \le nt} X_i = \sqrt{t} \frac{1}{\sqrt{nt}\sigma} \sum_{i \le nt} X_i \to \mathcal{N}(0, t).$$

Given $t < s$, we can represent

$$W_s^n = W_t^n + \frac{1}{\sqrt{n}\sigma} \sum_{nt < i \le ns} X_i$$

and since W_t^n and $W_s^n - W_t^n$ are independent, it should be obvious that the f.d. distributions of W_t^n converge to the f.d. distributions of the Brownian motion W_t. By Lemma 45, this identifies W_t as the unique possible limit of W_t^n and, if we can show that the sequence of laws $(\mathcal{L}(W_t^n))_{n\ge 1}$ is uniformly tight on $\big(C[0,\infty), d\big)$, Lemma 36 in Section 18 will imply that $W_t^n \to W_t$ weakly. Since $W_0^n = 0$, we only need to show equicontinuity (23.0.3). Let us write the modulus of continuity as

$$m^T(W^n,\delta) = \sup_{|t-s|\le\delta, t,s\in[0,T]} \Big| \frac{1}{\sqrt{n}\sigma} \sum_{ns<i\le nt} X_i \Big| \le \max_{0\le k\le nT, 0<j\le n\delta} \Big| \frac{1}{\sqrt{n}\sigma} \sum_{k<i\le k+j} X_i \Big|.$$

If instead of maximizing over all $0 \le k \le nT$, we maximize over $k = ln\delta$ for $0 \le l \le m-1$, $m := T/\delta$, i.e. in increments of $n\delta$, then it is easy to check that the maximum will decrease by at most a factor of 3, because the second maximum over $0 < j \le n\delta$ is taken over intervals of the same size $n\delta$. As a consequence, if $m^T(W^n,\delta) > \varepsilon$ then one of the events

$$\Big\{ \max_{0<j\le n\delta} \Big| \frac{1}{\sqrt{n}\sigma} \sum_{ln\delta<i\le ln\delta+j} X_i \Big| > \frac{\varepsilon}{3} \Big\}$$

must occur for some $0 \le l \le m-1$. Since the number of these events is $m = T/\delta$,

$$\mathbb{P}\Big(m^T(W^n,\delta) > \varepsilon\Big) \le m\,\mathbb{P}\Big(\max_{0<j\le n\delta} \Big| \frac{1}{\sqrt{n}\sigma} \sum_{0<i\le j} X_i \Big| > \frac{\varepsilon}{3} \Big). \tag{23.0.4}$$

Kolmogorov's inequality, Theorem 11 in Section 6, implies that if $S_n = X_1 + \cdots + X_n$ and

$$\max_{0<j\le n} \mathbb{P}(|S_n - S_j| > \alpha) \le p < 1$$

then

$$\mathbb{P}\Big(\max_{0<j\le n} |S_j| > 2\alpha\Big) \le \frac{1}{1-p} \mathbb{P}(|S_n| > \alpha).$$

If we take $\alpha = \varepsilon\sqrt{n}\sigma/6$ then, by Chebyshev's inequality,

$$\mathbb{P}\Big(\Big| \sum_{j<i\le n\delta} X_i \Big| > \frac{1}{6}\varepsilon\sqrt{n}\sigma \Big) \le \frac{6^2\delta n\sigma^2}{\varepsilon^2 n\sigma^2} = 36\delta\varepsilon^{-2}$$

and, therefore, if $36\delta\varepsilon^{-2} < 1$,

$$\mathbb{P}\Big(\max_{0<j\le n\delta} \Big| \sum_{0<i\le j} X_i \Big| > \frac{1}{3}\varepsilon\sqrt{n}\sigma \Big) \le \big(1 - 36\delta\varepsilon^{-2}\big)^{-1} \mathbb{P}\Big(\Big| \sum_{0<i\le n\delta} X_i \Big| > \frac{\varepsilon}{6}\sqrt{n}\sigma \Big).$$

Finally, using (23.0.4) and the central limit theorem,

$$\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}(m^T(W_t^n,\delta) > \varepsilon) &\le m\big(1 - 36\delta\varepsilon^{-2}\big)^{-1} \limsup_{n\to\infty} \mathbb{P}\Big(\Big| \sum_{0<i\le n\delta} X_i \Big| > \frac{\varepsilon}{6}\sqrt{n}\sigma \Big) \\
&= m\big(1 - 36\delta\varepsilon^{-2}\big)^{-1} 2\mathcal{N}(0,1)\Big(\frac{\varepsilon}{6\sqrt{\delta}}, \infty\Big) \\
&\le 2T\delta^{-1}\big(1 - 36\delta\varepsilon^{-2}\big)^{-1} \exp\Big(-\frac{1}{2}\frac{\varepsilon^2}{6^2\delta}\Big) \to 0
\end{aligned}$$

as $\delta \to 0$. This proves that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}(m^T(W^n,\delta) > \varepsilon) = 0,$$

for all $T > 0$ and $\varepsilon > 0$ and, thus, $W_t^n \to W_t$ weakly in $(C[0,\infty), d)$.

∎